An adverbial approach for the formal specification of topological constraints involving regions with broad boundaries

L. Bejaoui, F. Pinet, Michel Schneider, Y. Bedard

An Adverbial Approach for the Formal Specification of Topological Constraints Involving Regions with Broad Boundaries

Lotfi Bejaoui[1,2,3,4], François Pinet[3], Michel Schneider[3,4], and Yvan Bédard[1,2]

[1] Centre for Research in Geomatics (CRG), Laval University, Quebec (QC) Canada
lotfi.bejaoui.1@ulaval.ca
[2] Industrial Research Chair in Geospatial Databases for Decision Support, Laval University, Quebec (QC), Canada
yvan.bedard@scg.ulaval.ca
[3] Cemagref-Clermont-Ferrand, France
francois.pinet@cemagref.fr
[4] Dept. Computer Sciences, Blaise-Pascal University, Clermont-Ferrand, France
michel.schneider@isima.fr

Abstract. Topological integrity constraints control the topological properties of spatial objects and the validity of their topological relationships in spatial databases. These constraints can be specified by using formal languages such as the spatial extension of the *Object Constraint Language* (OCL). Spatial OCL allows the expression of topological constraints involving crisp spatial objects. However, topological constraints involving spatial objects with vague shapes (e.g., *regions with broad boundaries*) are not supported by this language. Shape vagueness requires using appropriate topological operators (e.g., *strongly Disjoint, fairly Meet*) to specify valid relations between these objects; otherwise, the constraints cannot be respected. This paper addresses the problem of the lack of terminology to express topological constraints involving regions with broad boundaries. We propose an extension of *Spatial OCL* based on a geometric model for objects with vague shapes and an adverbial approach for topological relations between regions with broad boundaries. This extension of *Spatial OCL* is then tested on an agricultural database.

1 Introduction

Internal spatial data quality is judged by several components, including completeness and logical consistency [14, 24]. Logical consistency is defined as the number of features, relationships, or attributes that have been correctly encoded in accordance with the integrity constraints [7, 19, 21] for the feature data specification [14]. Integrity constraints are defined at the conceptual level through specific tools [1]. In spatial databases, additional integrity constraints are required to control topological properties of geometries (e.g., *line simplicity*), semantic aspects (e.g., *a house has one level at least*), and topological relations (e.g., *agricultural spread parcels should be disjoint or adjacent)* in addition to basic constraints (e.g., *domain constraints*) [13, 22]. In this paper, we are interested in integrity constraints involving topological relations in transactional databases.

Q. Li et al. (Eds.): ER 2008, LNCS 5231, pp. 383–396, 2008.

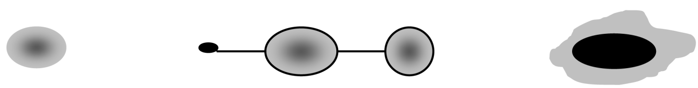
Formal specification of topological integrity constraints requires using an unambiguous formal language adapted to spatial databases. A spatial database-oriented language should allow the specification of both alphanumeric and spatial constraints [10, 16]. Currently, an extension of the Object Constraint Language (OCL) called Spatial OCL [10, 16] allows formal expression of spatial integrity constraints. Spatial OCL is based on the 9-Intersection model [11]. OCL provides a framework to define integrity constraints on classes' attributes or to differentiate between classes by using the *navigation* concept. This language has several advantages. First, it is easier to write an OCL constraint than its corresponding SQL query. Second, it is considered a subset of UML and based on the object-oriented paradigm commonly used in the software engineering domain. However, Spatial OCL cannot define topological constraints involving objects with vague shapes such as regions with broad boundaries [3, 4, 6, 8, 9, 17, 23]. These objects cannot be presented through crisp shapes [25] and therefore their topological relations cannot be identified by applying a spatial model for crisp objects such as the 9-Intersection model [11] or the CBM method [5]. For example, an integrity constraint may state that "a pollution zone *A* should not *overlap* a pollution zone *B*." The topological operator *overlap* cannot have the same definition as in the 9-Intersection model [11], because pollution zones can be viewed as regions with broad boundaries. They are not composed of the same topological invariants as crisp regions (they have broad boundaries instead linear ones) [18, 26]. Then, these regions with broad boundaries can overlap each other with different strengths: *weakly*, *fairly*, *strongly*, or *completely*. A classification of integrity constraints involving objects with vague shapes has been proposed in [2]. In this paper, we address the problem of the lack of terminology in Spatial OCL [10, 16] to express topological constraints involving objects with vague shapes. The main objective of this paper is to extend Spatial OCL in order to support topological constraints for regions with broad boundaries. We aim to extend the meta-model of Spatial OCL by proposing new types for objects with vague shapes and new topological operators adapted to regions with broad boundaries.

The paper is organized as follows. In section 2, we briefly review the notion of objects with vague shapes. Then, we present a spatial model for regions with broad boundaries and qualitative identification of their topological relations according to the *Qualitative Min-Max* (*QMM*) model presented in [3]. In section 3, we review related works on the specification of topological constraints, especially the approach using Spatial OCL [10, 16]. In section 4, we present our extension of Spatial OCL in order to formally express topological relations between regions with broad boundaries by using the QMM model [3]. Section 5 presents an example of a spatial database storing information about agricultural spreading activities. Some spatial objects stored in this database such as spread parcels have vague shapes, and therefore their topological constraints are expressed by using the extension of Spatial OCL. Section 6 presents the conclusions of this work.

2 Objects with Vague Shapes

2.1 Categorization of Spatial Objects with Vague Shapes

According to [12, 15], *shape vagueness* refers to the difficulty of distinguishing the shape of one object from its neighborhood. It is an intrinsic property of an object that has a spatial extent in a known position but does not have a well-defined shape (e.g., a pollution zone, a lake, a forest stand, etc.) [12]. We distinguish three basic types of *spatial objects with vague shapes*: *broad points*, *lines with vague shapes* (i.e., *lines with broad boundaries*, *lines with broad interiors* or *broad lines*), and *regions with broad boundaries*. Figure 1 shows an example of each one of these types of objects. A region has a vague shape when it is surrounded by a broad boundary instead of a sharp one (Figure 1c); we refer to these as *regions with broad boundaries* (e.g., a pollution zone). A line has a vague shape when its boundary (endpoints) and/or its interior are broad (Figure 1b; e.g., the itinerary of an historic explorer). For lines, we make a distinction between *broad interior* and *broad boundary* as we consider them specializations of *linear shape vagueness*. This distinction is also useful for points because a point does not have a boundary; it is only composed of an interior. A point's shape corresponds to the elementary space portion, which refers to its interior (Figure 1a). A broad point arises when there is a difficulty to distinguish the punctual object from its neighborhood (e.g., a mountain peak).



(a) A broad point (b) A line with a vague shape (c) A region with a broad boundary

Fig. 1. Examples of objects with vague shapes

Figure 2 shows our general categorization of objects with vague shapes. Three types of objects with vague shapes are specified: region with a broad boundary, line with a vague shape and broad point. Shape vagueness for lines can be a property of their boundaries (endpoints) and/or interiors. A line has a broad boundary when one of the endpoints at least is broad. A line with a vague shape can also correspond to a line where the interior is partially or completely broad; we speak about lines with broad interiors. The constraint *Overlap* means that a line may combine different types of shape vagueness. A line can have a broad boundary and a broad interior at the same time. Finally, a line can be completely broad when there is a difficulty to distinguish each point of the line from its neighborhood.

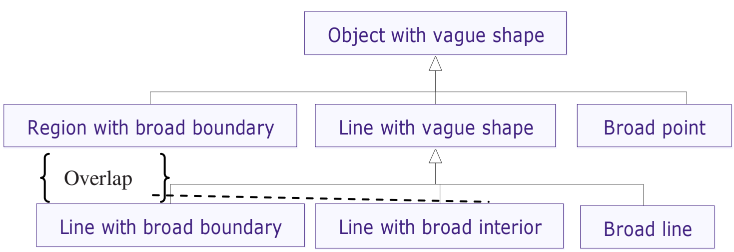
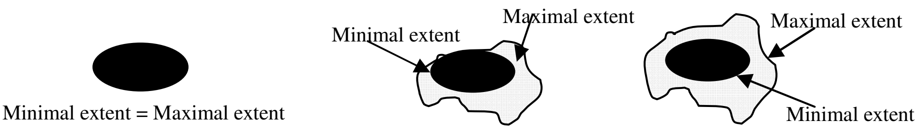


Fig. 2. Categorization of objects with vague shapes

2.2 Regions with Broad Boundaries and Their Topological Relations

In this paper, we define a region with a broad boundary according to the *QMM* model [3]. A region with a broad boundary is then composed by two crisp subregions: (1) *a maximal extent* A_{max} (i.e., the representation of the region when the boundary is considered as far as possible) and (2) *a minimal extent* A_{min} (i.e., the representation of the region when the boundary is considered as close as possible). These two extents should are related by one of the following topological relations: *Equal*[1](A_{min}, A_{max}) *or* *Contains*(A_{min}, A_{max}) *or* *Covers*(A_{min}, A_{max}) (Figure 3). The broad boundary refers to the difference between these two extents. This difference may include area everywhere around the minimal extent (i.e., regions with completely broad boundaries), may include area in some locations but not others around the minimal extent (i.e., regions with partially broad boundaries) or empty everywhere around the minimal extent (i.e., regions with no broad boundaries, or crisp regions). In Figure 3b, we present an example of a region with a partially broad boundary. The boundary is partially broad because the difference between the maximal extent and the minimal one is empty in some locations. Figures 3a and 3c, represent an example of a crisp region and another one of a region with a completely broad boundary, respectively.



(a) A crisp region (b) A region with a partially broad boundary (c) A region with a completely broad boundary

Fig. 3. Regions with broad boundaries

In order to specify topological relations between two regions with broad boundaries, we apply the 9-Intersection model [11] to identify the subrelations between the minimal and maximal extents of regions involved [3]. These subrelations are described through a 4-Intersection matrix including the values $R_1(A_{min}, B_{min})$, $R_2(A_{min}, B_{max})$, $R_3(A_{max}, B_{min})$, and $R_4(A_{max}, B_{max})$, where A and B are two regions with broad boundaries. Each cell of the 4-Intersection matrix receives one of the eight possible topological relations between two simple crisp regions (i.e., *Disjoint*, *Overlap*, *Meet*, *Equal*, *Contains*, *Inside*, *Covers*, *Covered by*). The 4-Intersection matrix corresponds to the following representation:

$$\begin{array}{cc} & \begin{array}{cc} B_{min} & \quad B_{max} \end{array} \\ \begin{array}{c} A_{min} \\ \\ A_{max} \end{array} & \begin{bmatrix} R_1(A_{min}, B_{min}) & R_2(A_{min}, B_{max}) \\ & \\ R_3(A_{max}, B_{min}) & R_4(A_{max}, B_{max}) \end{bmatrix} \end{array}$$

By considering eight possible values in a matrix' cells, $8^4 = 4096$ matrices can be distinguished. However, the definition of regions with broad boundaries specifies that only

[1] The spatial relations (i.e., *Equal*, *Contains*, *Covers*) used in this definition are those defined in (Egenhofer and Herring, 1990).

three relations between minimal and maximal extents are possible: *Equal*(A_{max}, A_{min}), *Contains*(A_{max}, A_{min}), or *Covers*(A_{max}, A_{min}). Thus, the contents of the matrix's cells are not mutually independent. For example, if the maximal extents are disjointed, it is inconsistent to have an *Overlap* relation between the minimal extents. By studying the possible consistency of matrices describing topological relations, we deducted that only 242 topological relations are possible between two simple regions with broad boundaries [3]. With regards to the content of a matrix, a topological relation can be classified into different clusters. Since eight values are possible in each cell of the 4-Intersection matrix, eight basic clusters can be distinguished: *DISJOINT*, *CONTAINS*, *COVERS*, *COVEREDBY*, *INSIDE*, *MEET*, *OVERLAP,* and *EQUAL*. In [3], we used four adverbs in order to qualify the membership of one relation to the clusters involved: *weakly* (only one of the matrix's cells has the same name as the cluster), *fairly* (two of the matrix's cells have the same name as the cluster), *strongly* (three of the matrix's cells have the same name as the cluster), and *completely* (all of the matrix's cells have the same value). Then, we distinguish for each basic cluster four subclusters which refer to the four levels of membership specified above: *weakly*, *fairly*, *strongly* and *completely*. Figure 4 presents some relations which belong to different subclusters of *CONTAINS* and *DISJOINT* clusters according to the contents of their respective matrices.

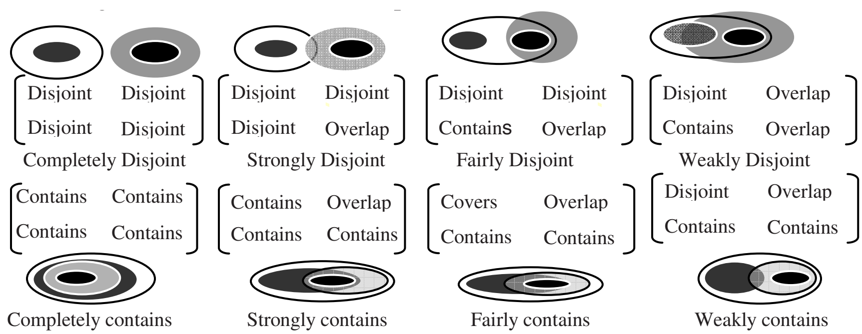


Fig. 4. Qualification of a topological relation between two regions with broad boundaries

In Figure 4, the fourth topological relation of the first line belongs with different strengths to the following clusters: *weakly* to *DISJOINT cluster*, *weakly* to *CONTAINS cluster,* and *fairly* to *OVERLAP cluster*. Hereafter, we integrate this adverbial approach into the object constraint language Spatial OCL.

3 Specification of Topological Constraints in Spatial Databases

3.1 Integrity Constraints in Spatial Databases

In spatial databases, additional integrity constraints are required to insure consistency of spatial objects [7, 21, 22]. In this paper, we are interested in topological constraints. These constraints control the validity of topological relations between spatial objects. We study formal expression of these constraints for regions with broad boundaries by using an extension of Spatial OCL.

3.2 OCL and Spatial OCL

OCL is a formal language that can be used to model invariants on UML models [16, 20]. These invariants can correspond to the integrity constraints of a database. Integrity constraints are defined in an UML class diagram. They correspond to conditions that must be satisfied for all instances of a class at any time. The class ruled by the constraint is called *context*. The principle of *navigation* consists in specifying integrity constraints which involve objects of different classes by using their associations. The following constraint specifies that the distance of an agricultural spread parcel from the closest lake must be greater than 100 meters:

```
Context Spreading_Parcel inv:
self.distance_lake > 100
```

In order to define spatial integrity constraints, Duboisset et al. [10] and Pinet et al. [16] proposed an extension of OCL's meta–model. This extension consists in adding geographic basic types (i.e., *point*, *line*, and *region*) to the meta-model of OCL (Figure 5). Moreover, topological relations can be expressed through Spatial OCL by using eight new topological operators added to the language: *overlaps*, *contains*, *is inside*, *are adjacent*, *covers*, *is covered by*, *are disjoint*, and *are equal*. These operators correspond to the topological relations defined in the 9-Intersection model [11]. For example, the topological constraint “buildings and roads *should not overlap each other*” is specified as follows:

```
Context road inv:
Building.allInstances→forAll(b|Self.geometry→aredisjoint(b)or
self.geometry→areAdjacent(b))
```

Additional OCL extensions are required to deal with topological constraints for regions with broad boundaries. For example, how can we express a topological constraint which specifies that "two pollution zones should be *completely disjoint* or *fairly meet* each other"? We need more tolerant topological operators than those currently used in Spatial OCL. Hereafter, we propose an extension of the Spatial OCL in order to support the formal expression of topological constraints between regions with broad boundaries. We call this extension Adverbial Spatial OCL for *Objects with vague shapes* ($AOCL_{OVS}$ for short). $AOCL_{OVS}$ is based on the *QMM* spatial model [3] and It consists of integrating a set of keywords of Spatial OCL in order to express the strength of topological relations specified in a constraint.

4 Adverbial Spatial OCL for Objects with Vague Shapes ($AOCL_{OVS}$)

In Spatial OCL [10, 16], geographic types are generalized through an abstract type called *BasicGeoType*. *BasicGeoType* allows definition of constraints on spatial attributes called *geometry*. Each value of *geometry* attribute value is a bag of elements; the type of each element is *BasicGeoType*. In order to consider vague shapes, we propose two abstract subclasses of geometries generalized by *BasicGeoType*: a type for *Objects with vague shapes* (*OVSType*) and another one for *Objects with Crisp Shapes* (*OCSType*). *OVSType* is a generalization of three basic types of objects with vague shapes: *broad point*, *line with a vague shape* and *region with a broad boundary*. A region with a broad boundary is composed by two crisp polygons (i.e., this relation is expressed through an aggregation between the type *Region with a broad boundary* and the type *Polygon*), which represent the minimal extent and maximal extent of the object, respectively. Figure 5 shows a general extension that covers three basic types of objects with vague shapes. In this paper, we focus on the topological constraints only for regions with broad boundaries.

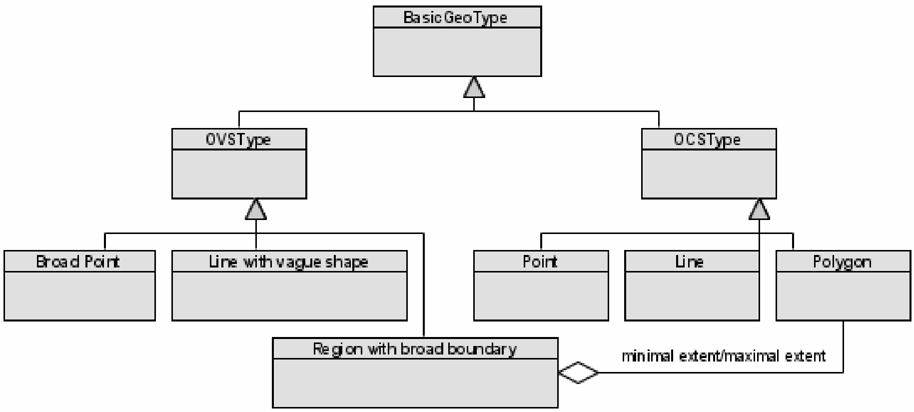


Fig. 5. Extension of the meta–model of Spatial OCL

The qualitative approach proposed in the *QMM* model [3] distinguishes 40 clusters (eight basic clusters and 32 subclusters) of topological relations between regions with broad boundaries (Section 2.2). Consequently, the proposed Spatial OCL extension introduces forty new topological operators adapted to regions with broad boundaries. These operators provide a qualitative evaluation of the strength of a topological relation. These operators can appear in OCL expressions when objects have the *OVSType* (Object with Vague Shape Type) and more precisely *Region with a broad boundary* type. A region with a broad boundary is considered valid when it verifies the next conditions:

1. Each one of the minimal extent and maximal extent verifies the closeness and connectedness conditions of a simple crisp region.
2. The minimal and maximal extents of a region with a broad boundary are related by one of the following topological relations: *Contains* (*max, min*), *Covers* (*max, min*), or *Equal* (*max, min*) (cf. section 2.2).

These last conditions are the *invariants* of the spatial model. We call these invariants *meta-constraints*, which control the validity of a region declared as a *Region with a broad boundary* (*RBB*).

5 Example in Agricultural Spreading Activities

Agricultural spreading activities consist of putting an organic substance *on* or *into* the soil in order to improve its agricultural productivity. In France, this activity is strictly controlled by public organizations, because substances used in spreading can be dangerous for ecological systems if they are not reasonably applied. The quantities and types of these substances depend on several criteria such as the parcel emplacement and soil type. For that, farmers should declare the areas to be spread and their references (i.e., they declare an *outline for the area to be spread*). Then, data about spreading activities are stored into a national spatial database. This database is accessed by a GIS-based tool available on the Web. The GIS-Based tool allows retrieving and updating of data describing spreading outlines declared by farmers. Farmers use the GIS-based tool to declare the areas of parcels before drawing their respective geometries on the screen through a GIS-based interface. The areas computed by the GIS tool for the drawn geometries of parcels are generally different from those declared by the farmer. Thus, a spread parcel has a theoretic geometry and an approximately drawn one. The difference between these two geometries corresponds to the broad boundary of a parcel. A spread parcel is a region with a broad boundary where the inner geometry corresponds to its minimal extent and the outer one corresponds to its maximal extent. The theoretic geometry is reconstructed from the drawn one by using the difference between the theoretic area and that of the drawn geometry. The area of a theoretic geometry should be equal to the drawn area.

Additionally, a spread parcel may be composed of one or several capacity zones that correspond to the parcel's subparts where the spreading is allowed with conditions (e.g., preserving the soil quality). Figure 6 shows an example of the theoretic geometry of a spread parcel (P_{Theo}), the drawn geometry of the same spread parcel (P_{Dr}), the theoretic geometry of a capacity zone Z1 ($Z1_{PTheoc}$), and the drawn geometry of a capacity zone Z1 ($Z1_{PDr}$). In this paper, we present a part of the conceptual schema of our spatial database (Figure 7). The class *Parcel* refers to an agricultural parcel contained by a spreading perimeter. A parcel is described by an identifier, a declared area, an area computed from the drawn geometry (*Draw_area*), and geometry with a vague shape composed by the drawn geometry and the theoretic one. Capacity zones are also defined as regions with broad boundaries. Finally, a spreading perimeter is a global area containing one or several spread parcels. Figure 7 presents a part of the class diagram of the spatial database storing data about agricultural spreading activities.

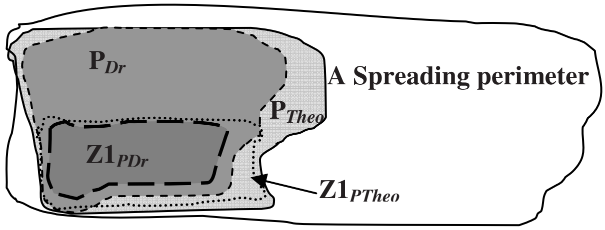


Fig. 6. An example of spatial data stored in the spreading agricultural database

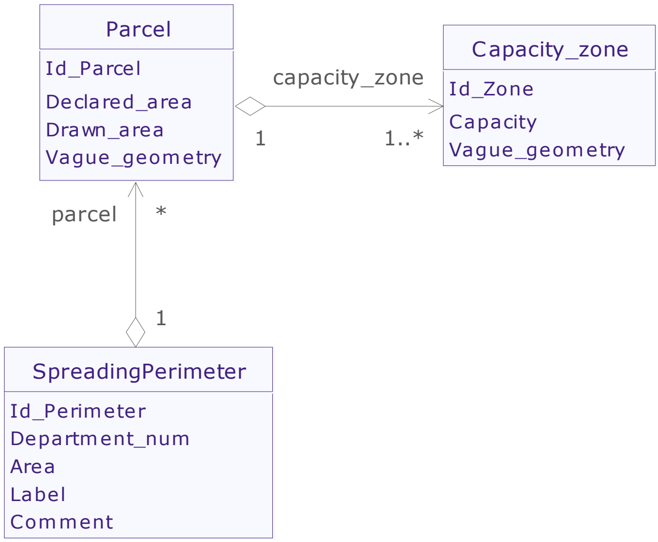


Fig. 7. Class diagram of the agricultural spreading database

5.1 Formal Expression of Constraints

The constraints presented below are expressed by using the $AOCL_{OVS}$ and they principally concern the spread parcels and their capacity zones. In this section, the *maximal extent* of one spread parcel refers to the theoretic geometry whether it covers or contains the drawn area, which is the minimal extent of the region in this case. In the same way, the drawn geometry refers to the maximal extent whether it covers or contains the theoretic geometry, which refers to the minimal extent in this second case. The minimal extent refers to the intersection of the theoretic geometry and the drawn one if they overlap each other. In this last case, the maximal extent refers to the union of the theoretic geometry and the drawn one.

Constraint 1: In a spreading outline, the parcels declared by farmers should be disjointed or meet each other. In the same way, the drawn geometries of these parcels, which have been manually drawn through a GIS-based tool, should also verify one of the topological relations: *Disjoint* or *Meet*. In our database, a parcel is an object with a vague shape, because a broad boundary results from the difference between the theoretic and drawn geometry. The topological relation between two parcels is valid when it belongs to one of the following subclusters: *completely Disjoint* (i.e., when both minimal and maximal extents are disjointed), *completely Meet* (i.e., when both minimal and maximal extents meet each other), *strongly Disjoint* and *weakly Meet* (i.e., when maximal extents meet each other but minimal extents are disjointed), or *fairly Disjoint* and *fairly Meet* (i.e., when maximal extents meet each other, minimal extents are disjointed, and one of the minimal extents meets one of the maximal extents):

```
Context Parcel inv:
Parcel.allInstances → forAll (b| self<>b implies
self.vague_geo→completely Meet(b.vague_geo) or
self.vague_geo→completely Disjoint(b.vague_geo) or
(self.vague_geo→strongly Disjoint(b.vague_) and self.vague_geo→weakly
Meet(b.vague_geo)) or (self.vague_geo→fairly Disjoint(b.vague_geo) and
self.vague_geo→fairly Meet(b.vague_geo)))
```

Constraint 2: A spread parcel is composed by one or several capacity zones. A capacity zone is inside, and covered by or equal to the drawn geometry of the parcel involved. The same relations should be respected between respective theoretic geometries of a parcel and each of its capacity zones. Indeed, the topological relation between a parcel and each of its capacity zones (both represented as regions with broad boundaries) is valid if it belongs to one of the following subclusters: *completely Contains*, *completely Covers*, *strongly Contains* and *weakly Covers*, *strongly Contains* and *weakly Overlap*, *fairly Contains* and *fairly Covers*, *fairly Contains* and *weakly Covers* and *weakly Overlap*, *strongly Covers* and *weakly Contains*, *fairly Contains* and *fairly Covers*, or *strongly Covers* and *weakly Overlap*:

```
Context Parcel inv:
self.vague_geo→ forAll (b| self.capacity_zone.vague_geo→ exists(d|
(b.vague_geo→completely Contains(d.vague_geo)) or
(b.vague_geo→completely Covers(d.vague_geo)) or (b.vague_geo→strongly
Contains(d.vague_geo) and b.vague_geo→weakly Covers(d.vague_geo)) or
(b.vague_geo→strongly Contains(d.vague_geo) and b.vague_geo→weakly
Overlap(d.vague_geo)) or (b.vague_geo→fairly Contains(d.vague_geo) and
b.vague_geo→fairly Covers(d.vague_geo)) or (b.vague_geo→fairly
Contains(d.vague_geo) and b.vague_geo→weakly Covers(d.vague_geo) and
b.vague_geo→weakly Overlap(d. vague_geo)) or (b.vague_geo→strongly
Covers(d.vague_geo) and b.vague_geo→weakly Contains(d.vague_geo)) or
(b.vague_geo→fairly Contains(d.vague_geo) and b.vague_geo→fairly
Covers(d.vague_geo)) or (b.vague_geo→strongly Covers(d.vague_geo) and
b.vague_geo→weakly Overlap(d.vague_geo))))
```

Constraint 3: Inside one spread parcel, two different capacity zones should verify one of the following specifications: *completely Disjoint*, *completely Meet*, (*strongly Disjoint and weakly Meet*) or (*fairly Disjoint and fairly Meet*).

```
Context Capacity_zone inv:
self.allInstances → forAll (a,b| a<>b and a.parcel=b.parcel implies a.
vague_geo→completely Meet(b.vague_geo) or a.vague_geo→completely
Disjoint(b.vague_geo) or (a.vague_geo→strongly Disjoint(b.vague_geo)
and a.vague_geo→weakly Meet(b.vague_geo)) or (a.vague_geo →fairly
Disjoint(b.vague_geo) and a.vague_geo→fairly Meet(b.vague_geo)));
```

Constraint 4: *P* is a spreading perimeter composed by *N* spread parcels. The sum of areas of minimal extents of spread parcels is less than or equal to the area of *P*. However, the sum of areas of maximal extents of spread parcels is greater than or equal to the declared area of *P*. The expression "*self.parcel.vague_geo.minimal_extent.area→sum()*" provides the sum of areas of minimal extents of parcels belonging to the spreading perimeter involved. In other words, this function makes the same thing for maximal extents of capacity zones in one spread parcel.

```
Context SpreadingPerimeter inv:

self.parcel.vague_geo.minimal_extent.area→sum() ≤ self.area and
self.parcel.vague_geo.maximal_extent.area→sum() ≥ self.area
```

5.2 Implementation of AOCL$_{OVS}$

In this work, OCL expressions can be automatically translated into SQL code by using a constraint editor called OCL2SQL initially developed by Tudresden University before to be extended by [10, 16], first for topological constraints for crisp regions and next, in the present paper, for regions with broad boundaries. Figure 8 shows the architecture of OCL2SQL application. It is a Java application in which constraints are defined in an UML class diagram stored in an *xmi* file. The constraints are written by using AOCL$_{OVS}$ specifications to be verified according the class diagram involved. OCL2SQL editor translates these constraints in SQL language, wherein new topological operators are defined as PL/SQL functions managed by the DBMS (Database Management System) Oracle. For example, the next constraint specifies that two pollution zones should be strongly disjointed. For this constraint we give the correspondent SQL code. The SQL script generated by OCL2SQL is then executed on the data stored in an Oracle spatial database in order to retrieve possible inconsistencies.

Constraint 5:

```
Context Pollution_zones inv:

Parcel.allInstances→forAll (b| self<>b implies self.vague_geo→
strongly Disjoint(b.vague_geo)
```

Oracle Spatial SQL:

```
select * from OV_Pollution_Zone SELF
where not (not exists ( (select PK6 from OV_ Pollution_Zone) minus
      select PK6 from OV_ Pollution_Zone SELF2  where (SELF.PK6 = SELF2.PK6) OR
         stronglyDisjoint((select PK4 from OV_VAGUE_GEO
            where PK4 in (select GEOMETRY_PK4 from
                                    OV_ Pollution_Zone where PK6 = SELF2.PK6)),
                                    (select PK4 from OV_VAGUE_GEO where PK4 in
                                    (select GEOMETRY_PK4 from OV_ Pollution_Zone
                            where PK6 = SELF2.PK6)) , OV_VAGUE_GEO)=0 ));
```

Figure 8 schematizes the architecture of the extension of OCL2SQL, which covers topological constraints involving regions with broad boundaries. This figure is adapted from [10].

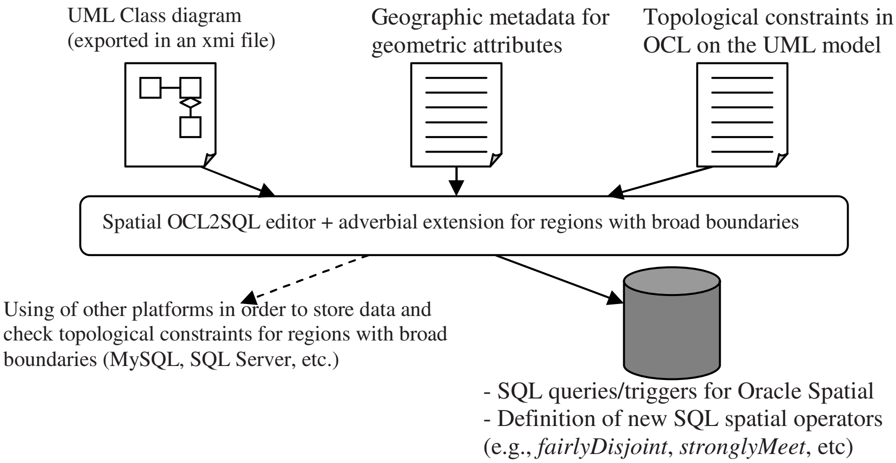


Fig. 8. Architecture of the application used to check the OCL constraints (this figure is adapted from [10])

6 Conclusion

Respecting topological constraints is an important aspect of internal spatial data quality. Topological constraints can be expressed through Spatial OCL [10, 16], which integrates the 9-Intersection model to specify topological relations. However, Spatial OCL lacks syntactical tools to express topological constraints for objects with vague shapes. In this paper, we addressed the problem of formal specification of topological constraints for objects with vague shapes and especially regions with broad boundaries. For that, we presented a spatial model for regions with broad boundaries, where topological relations are identified according to subrelations between their minimal and maximal extents [3]. Then, topological relations are qualitatively classified by exploring similarity between subrelations identified. Four adverbs are used to describe the strength of a topological relation between two regions with broad boundaries: *weakly*, *fairly*, *strongly*, or *completely*.

This paper makes three main contributions. First, the meta-model of Spatial OCL has been extended in order to consider new data types covering spatial objects with vague shapes. We proposed a new abstract type called *OVSType* (*Object with Vague Shape Type*), which can be specialized into *broad point*, *line with a vague shape*, and *region with a broad boundary*. Second, our adverbial approach for topological relations between regions with broad boundaries has been integrated into Spatial OCL. Forty new topological operators have been proposed as additional keywords of Spatial OCL in order to deal with topological constraints involving regions with broad boundaries. We have called this extension *Adverbial spatial OCL for Objects with Vague Shapes* (*AOCL$_{OVS}$* for short). Third, AOCL$_{OVS}$ has been integrated into the constraint editor OCL2SQL, which automatically generates Oracle Spatial SQL code of the topological constraints from their AOCL$_{OVS}$ expressions. This framework has been tested using a spatial database storing data about agricultural spreading activities. Some constraints have been specified for this database. These constraints principally involve spread parcels and their capacity zones presented as regions with broad boundaries.

In the future, we aim to extend this approach in two main directions. First, we will generalize our framework in order to specify topological relations involving different objects with vague shapes (i.e., *broad points*, *lines with vague shapes*, and *regions with broad boundaries*). Second, we will study the specification of topological constraints involving regions with vague complex shapes (e.g., regions with several kernels, regions composed by several subregions with broad boundaries).

References

1. Bédard, Y., Larrivée, S., Proulx, M.J., Nadeau, M.: Modeling Geospatial Databases with Plug-Ins for Visual Languages: A Pragmatic Approach and the Impacts of 16 Years of Research and Experimentations on Perceptory. In: Wang, S., et al. (eds.) ER Workshops 2004. LNCS, vol. 3289, pp. 17–30. Springer, Heidelberg (2004)
2. Bejaoui, L., Bédard, Y., Pinet, F., Salehi, M., Schneider, M.: Logical consistency for vague spatiotemporal objects and relations. In: The 5th International Symposium on Spatial Data Quality (ISSDQ 2007), Enschede, Netherlands (June 2007)
3. Bejaoui, L., Bédard, Y., Pinet, F., Schneider, M.: Qualified topological relations between objects with possibly vague shape. International Journal of Geographical Information Sciences (to appear, 2008)
4. Burrough, P.A., Frank, A.U.: Geographic Objects with Indeterminate Boundaries. Taylor & Francis, London (1996)
5. Clementini, E., Di Felice, P.: A Comparison of Methods for Representing Topological Relationships. Information Sciences 3, 149–178 (1995)
6. Clementini, E., Di Felici, P.: Approximate topological relations. International Journal of Approximate Reasoning 16, 173–204 (1997)
7. Cockcroft, S.: A Taxonomy of Spatial Data Integrity Constraints. Geoinformatica 1(4), 327–343 (1997)
8. Cohn, A.G., Gotts, N.M.: The 'egg-yolk' representation of regions with indeterminate boundaries. In: Burrough, P., Frank, A. (eds.) Proceedings of the GISDATA Specialist Meeting on Spatial Objects with Undetermined Boundaries, pp. 171–187. Taylor & Francis, Abington (1996)
9. Dilo, A.: Representation of and reasoning with vagueness in spatial information: A system for handling vague objects. PhD thesis, ITC, Netherlands, p. 187 (2006)
10. Duboisset, M., Pinet, F., Kang, M.A., Schneider, M.: Precise modeling and verification of topological integrity constraints in spatial databases: from an expressive power study to code generation principles. In: Delcambre, L.M.L., Kop, C., Mayr, H.C., Mylopoulos, J., Pastor, Ó. (eds.) ER 2005. LNCS, vol. 3716, pp. 465–482. Springer, Heidelberg (2005)
11. Egenhofer, M., Herring, J.: A mathematical framework for the definition of topological relations. In: Brassel, K., Kishimoto, H. (eds.) Proceedings of the Fourth International Symposium on Spatial Data Handling, Zurich, Switzerland, pp. 803–813 (1990)
12. Erwig, M., Schneider, M.: Vague regions. In: Scholl, M.O., Voisard, A. (eds.) SSD 1997. LNCS, vol. 1262, pp. 298–320. Springer, Heidelberg (1997)
13. Frank, A.U.: Tiers of ontology and consistency constraints in geographical information systems. Int. J. of Geographical Information Science 15(7), 667–678 (2001)

14. Guptill, S.C., Morrison, J.L.: Spatial data quality. In: Guptill, S.C., Morrison, J.L. (eds.) Elements of spatial data quality, Elsevier Science Inc., New York (1995)
15. Hazarika, S.M., Cohn, A.G.: A taxonomy for spatial vagueness, an alternative egg-yolk interpretation. In: Montello, D.R. (ed.) COSIT 2001. LNCS, vol. 2205, pp. 92–107. Springer, Heidelberg (2001)
16. Pinet, F., Duboisset, M., Soulignac, V.: Using UML and OCL to maintain the consistency of spatial data in environmental information systems. Environmental modelling & software 22(8), 1217–1220 (2007)
17. Reis, R., Egenhofer, M.J., Matos, J.: Topological relations using two models of uncertainty for lines. In: Proceeding of the 7th international Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Lisbon, Portugal, 5 - 7 July, pp. 286–295 (2006)
18. Rodriguez, A.: Inconsistency Issues in Spatial Databases. In: Bertossi, L., Hunter, A., Schaub, T. (eds.) Inconsistency Tolerance. LNCS, vol. 3300, pp. 237–269. Springer, Heidelberg (2005)
19. Salehi, M., Bédard, Y., Mir, A.M., Brodeur, J.: Classification of integrity constraints in spatiotemporal databases: toward building an integrity constraint specification language. International Journal of Geographical Information Science (submitted, 2007)
20. Schmid, B., Warmer, J., Clark, T.: Object Modeling with the OCL: the Relational Behind the Object Constraint Language, p. 281. Springer, Heidelberg (2002)
21. Servigne, S., Ubeda, T., Puricelli, A., Laurini, R.: A Methodology for Spatial Consistency Improvement of Geographic Databases. GeoInformatica 4(1), 7–34 (2000)
22. Souris, M.: Contraintes d'intégrité spatiales. In: Devillers, R., Jeansoulin, R. (eds.) Qualité de l'information géographique, Lavoisier, pp. 100–123 (2006)
23. Tang, T.: Spatial object modeling in fuzzy topological spaces: with applications to land cover change. PhD thesis, University of Twente (2004) ISBN 90-6164-220-5
24. Van Oort, P.: Spatial data quality: from description to application. In: Publication on Geodesy 60, Delft, December 2006, Geodetic Commission, Netherlands (2006)
25. Yazici, A., Zhu, Q., Sun, N.: Semantic data modeling of spatiotemporal database applications. Int. J. Intell. Syst., 881–904 (2001)
26. Zhan, F.B., Lin, H.: Overlay of Two Simple Polygons with Indeterminate Boundaries. Transactions in GIS 7(1), 67–81 (2003)